

The Sloan Digital Sky Survey QSO absorption line catalogue

Donald G. York¹, Daniel Vanden Berk², Gordon T. Richards³, Arlin P. S. Crotts⁴, Pushpa Khare⁵, James Lauroesch⁶, Martin Lemoine⁷, Scott Burles⁸, Mariangela Bernardi⁹, Francisco J. Castander¹⁰, Josh Frieman¹, Jon Loveday¹¹, Avery Meiksin¹², Robert Nichol¹³, David Schlegel¹⁴, Donald P. Schneider², Mark Subbarao¹, Chris Stoughton¹⁵, Alex Szalay¹⁶, Brian Yanny¹⁵, Yusra Alsayyad¹, Abhishek Kumar¹⁷, Britt Lundgren¹⁸, Natela Shanidze¹, Johnny Vanlandingham¹, Matthew Wood¹⁹, Britt Baugher²⁰, Jon Brinkmann²¹, Robert Brunner¹⁸, Masaaka Fukugita²², Patrick B. Hall²³, Timothy M. Heckman¹⁶, Lewis M. Hobbs¹, Craig J. Hogan²⁴, Lam Hui¹⁵, Edward B. Jenkins³, Daniel Kunstz¹⁶, Brice Menard²⁵, Osamu Nakamura²⁶, Jean M. Quashnock²⁷, Michael Stein¹, Aniruddha R. Thakar¹⁶, David Turnshek²⁸, Daniel E. Welty¹,
and the SDSS Collaboration

¹The University of Chicago, USA ²Penn State University, USA ³Princeton University, USA
⁴Columbia University, New York, NY, USA ⁵Utkal University, India ⁶Northwestern University,
USA ⁷Institut d'Astrophysique, Paris ⁸Massachusetts Institute of Technology, USA ⁹University
of Pennsylvania, USA ¹⁰Institut d'Estudis Espacials de Catalunya/CSIC, Barcelona, Spain
¹¹University of Sussex, UK ¹²University of Edinburgh, Scotland ¹³University of Portsmouth,
UK ¹⁴Lawrence Berkeley National Laboratory, USA ¹⁵Fermi National Accelerator Laboratory,
Batavia, IL, USA ¹⁶Johns Hopkins University ¹⁷New York University, USA ¹⁸University of
Illinois at Urbana-Champaign, USA ¹⁹University of California, Los Angeles, USA ²⁰University
of California, Santa Barbara, USA ²¹New Mexico State University, Las Cruces, NM, USA
²²University of Tokyo, Japan ²³York University, Toronto, Canada ²⁴University of Washington,
Seattle ²⁵Institute of Advanced Study, Princeton, NJ, USA ²⁶University of Nottingham, UK
²⁷Carthage College, Kenosha, WI, USA ²⁸University of Pittsburgh

Abstract. The spectra of the Sloan Digital Sky Survey (SDSS) are being used to construct a catalogue of QSO absorption lines, for use in studies of abundances, relevant radiation fields, number counts as a function of redshift, and other matters, including the evolution of these parameters. The catalogue includes intervening, associated, and BAL absorbers, in order to allow a clearer definition of the relationships between these three classes. We describe the motivation for and the data products of the project to build the SDSS QSO absorption line catalogue.

1. Introduction

Absorption lines were found in the spectra of quasi-stellar objects as soon as the class of quasi-stellar sources was recognised (Sandage 1965). The earliest studies focused on strong absorbers in the short wavelength wing of the main QSO emission lines. Many of these first systems would be termed “associated systems” today. Bahcall, Peterson & Schmidt (1966) found one of the earliest systems that could be termed intervening and Bahcall (1968) established the reality of such systems. Bahcall & Spitzer (1969) presented plausibility arguments that foreground galaxies with extended halos could explain some of the intervening absorption lines. A third class of lines, the broad absorption lines

(BALs), were carefully defined by Weymann *et al.* (1979), in the context of intervening and associated systems.

There have been numerous studies that have confirmed the key characteristics of the intervening absorbers: they are normally multi-component systems of absorption lines represented by a wide range of ions at the same velocity: the components generally have abundances that are less than Solar abundances by at least a factor of 10 (Prochaska *et al.* 2003; Khare *et al.* 2004; Kulkarni *et al.* 2005a). The same anomalous appearance of C IV and O VI is noted in the QSOALS as in our own Galaxy (Pettini & West 1982; Sembach *et al.* 2000; Wakker *et al.* 2003). There is some evidence for depletion of refractory elements relative to volatiles (e.g. Cr relative to Zn) as found for interstellar dust in our Galaxy (Hobbs *et al.* 2003, Fig. 2a). Continuum emission and/or line emission is sometimes seen from an associated galaxy, but it is often difficult to establish that the absorption lines are from the particular galaxy identified (Kulkarni *et al.* 2005b, and other papers in this volume). Evidently, these systems can be powerful probes of abundances and dynamics in galaxies over a wide range of redshifts, but the problem of finding the galaxies responsible for the absorption (Kulkarni *et al.* 2005b) must be solved before the promise of such studies can be fully realised.

The associated systems, generally with apparent ejection velocities of $v/c < 0.017$, defined to be positive for gas ejected toward Earth, are more numerous than the candidate intervening systems, and are generally stronger, but have ionisation properties similar to those of the intervening systems, and, to some extent, of the BALs. It is not known if the associated systems are related in origin to either of the other two classes of lines.

The BALs include gas, sometimes resolved into components and sometimes not, that is ejected at high velocity (> 2000 km/sec, by definition, to over 50,000 km/sec) and which sometimes involves very small clouds that do not cover the source. Their relationship to associated systems and to intervening systems is not clear. BALs and intervening systems can be intermingled in redshift. It is not known if there is an evolutionary relationship between them, or between the associated systems and either of the other classes.

In this paper, we describe a new, large catalogue of QSO absorption line systems, (QSOALS), designed to allow the addressing of a number of issues that have arisen over the years but which cannot be addressed with small samples. In section 2, specific motivation for the survey is given. In section 3, the spectra of the Sloan Digital Sky Survey are described. Their use to measure absorption lines is taken up in section 4. Section 5 gives selected, preliminary results.

2. Motivation

A large survey of QSOALS is needed to fully define the three classes of systems noted above and their relationships to each other. Clues must lie in the abundances, number counts with redshift, clustering, and radiation field of the respective classes of lines, and in the evolution of these quantities. However, such studies require samples that are pure, well defined in sensitivity level, and adequate in size and completeness to detect such subtle effects as gravitational weak lensing. (Vanden Berk *et al.* 1996; Ménard & Péroux 2003; Scranton *et al.* 2005).

The SDSS spectra provide a resolving power just adequate to study abundances and to separate the important doublets, for the purpose of statistical studies. For each object, SDSS provides information on extinction (Richards *et al.* 2003; Reichard *et al.* 2003a; Khare *et al.* 2005), high precision photometry of the background QSO, well defined limits for equivalent widths in regions where no absorbers are detected, and moderately deep images in five colours to examine the region near the absorber lines of sight for individual

galaxies and clusters. Finally, the survey offers enough spectra to perform statistical studies comparable in precision to those done heretofore for galaxies. Churchill (2005) includes a partial list of the much smaller samples encountered in previous surveys.

3. The SDSS spectra

This new cataloguing project was undertaken with the spectra of the Sloan Digital Sky Survey (SDSS) (York *et al.* 2000). Typically, the survey produces well defined systems accounting for all absorption lines down to the QSO Lyman alpha emission line for a given QSO. The equivalent width (absorber rest frame) sensitivity is 50-200 mÅ. The ultimate survey will be based on about 100,000 QSOs and contain at least that many absorption line systems.

The SDSS targets QSOs based on colour selection of objects to an i-magnitude of 20 based on 5-filter imaging scans of the Northern Galactic cap of the Milky Way, minimising the effects of Galactic reddening (Gunn *et al.* 1998). Using a precise, absolute flux calibration (Hogg *et al.* 2001; Smith *et al.* 2002; Ivezić *et al.* 2004) of the object images in the SDSS filter system (Fukugita *et al.* 1996), QSOs can be picked out and their redshifts can be accurately estimated, using only the colour data (Richards *et al.* 2004; Weinstein *et al.* 2004). Repeated testing of the technique shows that the end to end completeness of the QSO survey itself is about 90% (Vanden Berk *et al.* 2005).

The SDSS candidate QSOs are observed with a 640 fibre spectrographic system. Holes for the fibres are drilled into plates using absolute coordinates determined in the imaging reduction to better than 0.1 arc-sec (Pier *et al.* 2003). About 100 QSO candidates are chosen for each plate, according to the target selection prescription in Richards *et al.* (2002a). Developing an evolving plan for the plates, each of which contains QSOs, galaxies, and stars that are targeted for various reasons, and the overlap of the plates on particular regions of the sky, are described by Blanton *et al.* (2003). The resulting spectra have a resolving power of about 1800 (140 km/sec) with two-pixel sampling. The S/N is set by the procedure that 15 minute exposures are recorded until the sum of spectra of a single QSO at a g-magnitude of 20.2 gives $(S/N)^2 \sim 7$. Thus, the signal-to-noise for detection of absorption lines varies with magnitude. The spectral coverage is from 3800 Å to 9000 Å. The pipeline reduction of the spectra outputs the S/N at each data point, flags known defects, identifies the QSO emission lines, and derives a redshift for the QSO. The redshift takes account of the fact that the actual centres of the C IV and Mg II emission lines, for example, can have very different redshifts (Tytler & Fan 1992; Richards *et al.* 2002b).

Each spectrum is then examined by eye and placed in an official list of SDSS QSOs (one line broader than 1000 km/sec) with no restriction that the object be a point source image. These QSOs are published in lists by data release (DR) number: see Schneider *et al.* (2003) for DR1 and Schneider *et al.* (2005) for DR3. Selection of objects as candidate QSOs for taking SDSS spectra occurs via a strict set of rules relating to colour space for the purpose of reducing the number of false positives (white dwarfs, A stars) (Richards *et al.* 2002a). Fibres cannot be placed within 55 arc-sec of each other: plate overlaps are necessary for obtaining spectra of close pairs.

SDSS spectra of QSOs with absorption lines have been published in conjunction with broad absorption line QSOs (Hall *et al.* 2002a; Menou *et al.* 2001; Tolea, Krolik & Tsvetanov 2002; Reichard *et al.* 2003) and lensed QSOs (Inada *et al.* 2003; Hall *et al.* 2002b).

Specific SDSS data products are described by Stoughton *et al.* (2001) and Abazajian *et al.* (2003, 2005). Spectra, colours, and finding charts for all objects in DR3 are on the SDSS website.

4. The SDSS absorption line catalogue

A separate pipeline has been developed (York *et al.* 2001; York *et al.* 2005a,b, in preparation) to isolate absorption lines and to identify systems of lines. The systems are organised into a form modelled on the catalogue of York *et al.* (1991). The completeness of the current QSOALS catalogue is adequate for selecting definite low-ionisation or high-ionisation systems and for computing composite spectra of QSOALS with high signal-to-noise, as well as for preliminary statistical tests of the behaviour of the QSOALS.

The main absorption line catalogue is derived from a list of identified lines that are organised into systems. The system construction is keyed to finding Mg II or C IV doublets. (Extensive studies have been done to confirm that no systems are missed with this technique, unless the doublets fall out of range of the spectrographs. We hope to generalise this aspect of the search.)

All official SDSS QSOs are passed through our absorption line pipeline. The spectra are normalised using an automatic continuum-tracing procedure. A list of significant features in absorption (four-sigma detection threshold) is defined, using a program evolved from Vanden Berk *et al.* (1999), Richards (2001), and references therein. The equivalent widths, line widths, redshifts, and errors are measured automatically, using a profile fit to get the equivalent width. Systems are then automatically defined by searching for lines with identical redshifts. Grade A systems are defined as systems with at least 4 absorption lines, of four-sigma or greater significance, at the same redshift, none of which are blended with lines from other systems or in the Lyman alpha forest, or affected by artifacts of the observation. The agreement with the predicted wavelength of a given line for a given system is allowed to vary according to the signal-to-noise ratio of a given spectrum. Multiple lines typically agree to better than 300 km/sec, but the agreement may be as poor as 1000 km/sec in lower quality spectra. The mean redshift for the system so defined is determined, and the data for each line, including the single line redshifts, are stored in what is termed a catalogue file. Systematic differences between, for instance, Mg II and Fe II lines may be as low as 30 km/sec, while C IV may be different by a few hundred kilometres per second, an effect that is not caused by measuring error but by real shifts in velocity between regions of different ionisations.

Grade B, C, and D systems are also defined. The definition of such systems differs from those of sample A only in the number of identified lines in the system: three are needed for class B, two for class C, and one for class D. Class D includes, for example, systems that for various reasons can only be Mg II absorbers, but for which one of the lines is not seen owing to a blend with an uncorrected night-sky emission line.

The catalogue takes the form of labelled spectra; lists of identified lines of systems found; lists of all unidentified lines in the spectra; lists of any QSOs or galaxies within one arc-minute of a particular QSO with absorption lines; and various lists of the absorption line systems that can be generated to specification using an SQL interface to the absorption line web page (not yet public). An auxiliary file containing the normalised spectra is augmented with the equivalent-width limit for each resolution element, to allow total search ranges in z -space to be defined. Additional plots are provided that show the raw spectra, the fitted continuum, the positions of all lines that are either present or could be seen for all identified systems, and the positions of significant but unidentified absorption features. While grades are assigned based on lines of four-sigma significance,

three-sigma lines are listed and indicated on the plots. All lines of heavy elements that could be present are marked in the Lyman alpha forest: coincidences of lines in the forest with detected systems of heavy elements are noted in the system line lists, but most will be blends with Lyman alpha lines. The list of unidentified lines is essentially a list of Lyman alpha features with no counterparts among the lines of heavy elements, except for a few artifacts of the observations or reductions.

Broad absorption line systems (BALs) require special treatment because of their width. The continuum fitting is done with a 20-point smoothing algorithm that allows the discovery of most narrow absorption lines. When BALs occur that are much wider than 1000 km/sec, or when blends of narrow lines produce lines of similar width, the continuum is depressed by the absorption line itself and the equivalent-widths are lower limits: these are recognisable from the FWHM values for individual lines in the cat files. A second pass is then made through all the spectra, using a 50-point smoothing function. This pass ignores most narrow lines but properly detects and measures many BALs and blends. Only lines of C IV are treated in this way. Any systems found in this pass cause a flag to be generated that identifies the QSO as having a BAL feature in C IV. The SDSS project maintains a list of strong BALs, including those with BAL troughs too broad to be properly measured with our 50-point smoothing function. A flag is set if a given QSO is in the project BAL catalogue.

Using the measured lines, the system grades, and the flags, it is possible to construct lists of objects for further study. As an example, one can generate a list of grade A QSO absorption line systems with equivalent-widths above 0.3 Å, with $1 < z < 1.95$, in objects that contain no BAL features, have a particular limiting magnitude, and have an intrinsically weak line of a heavy element present (Zn II, Cr II, Mn II, etc.). Such lists were used to generate the samples used in the paper on the 2175 Å extinction bump and the extinction curves of QSOALS in this volume (Khare *et al.* 2005).

A log file has been set up for each QSO, into which particular problems with a given absorption line system can be noted, for the record. The log files are searchable, so the catalogue can be searched for persistent anomalies using string searches. An editing tool is being completed which will allow correction of misidentified lines, addition of systems missed by the automated search, or deletion of incorrect systems identified by the automated search.

The automated search currently identifies absorption line systems with 95% correctness (valid systems) and 95% completeness (there are systems that should have been identified but were not). There are several reasons for failure to derive a correct redshift or failure to find a system. These are related to, in the former case, blends, particularly of multiple C IV doublets short-ward of the C IV emission, possibly associated with insipid BALs. In the latter case, individual lines might be missed because they overlap with poorly corrected night sky emission or telluric absorption. Some of problems can be overcome with more sophisticated computer algorithms, but in a number of cases, human inspection is required. The catalogue will thus have different versions, consistent with generating a new catalogue for each SDSS data release. Each release includes all previously published spectra, re-reduced and therefore updated for any improvements in the basic SDSS spectro pipeline and the absorption line pipeline.

5. Preliminary results

In the process of testing the catalogue procedures, a number of undergraduate senior theses have been done by students at the University of Chicago. These include searches for damped Lyman alpha lines (T. Smith), quantification of the number of absorbers per unit

redshift (dM/dz) (B. Lundgren), a search for subtle BAL systems (Y. Alsayyad), a study of the relation between BALs and associated absorbers (Bin Choi), and a comparison of absorbers in the spectra of the components of gravitationally lensed QSOs and the intervening systems for which the associated galaxy is not seen directly (N. Shandize). We have used the catalogue measurements of equivalent-widths to pick candidates for studies of abundances in QSOALS at intermediate redshift (Khare *et al.* 2004). Extinction attributable to QSOALS have been studied by Khare *et al.* (2005). Other studies, to determine dM/dz for the main absorption lines (Si IV, C IV, Al III, Al II, Fe II, Mg II, and Mg I) are underway (York *et al.* 2005a, in preparation). A study of clustering of QSO absorbers is underway (Vanden Berk *et al.*, 2005, in preparation). Studies of the average properties of QSOALS in DR1, following the work of Nestor *et al.* (2003), who used EDR spectra of the SDSS, are also underway (York *et al.* 2005b, in preparation).

We have derived dM/dz for Mg II, Al II, and Si II, using lines at 2795, 1670, and 1526 Å, respectively, allowing us to trace the dM/dz for first ions over the range from $z = 0.4$ to $z = 4.4$, an unprecedented range for such a study. The function rises from $z = 0.4$ to $z = 2$, then levels off or drops slightly, showing that the space density of the absorbers is evolving at the higher redshifts. A similar study of C IV, Si IV, and Al III shows that despite the difference in ionisation potential for the three species, the curves in dM/dz all flatten out above $z = 2$, again, showing that there is evolution in the absorber number counts when one looks at high enough z . For all ions, the detections are significantly lower than would be expected for a population of unevolving objects.

6. Concluding Remarks

We hope to release the data set for public use following the completion of verification procedures now underway, to assure the quantifiable completeness of the catalogue. Our own work will be aimed at addressing the questions involving the nature of the three types of absorbers (intervening, associated, and BAL) and their relationships to each other. Such a goal involves studies of the properties of each class with regard to clustering; abundances, both absolute and as a function of redshift; and the large scale ionisation field, as compared to indicators from other kinds of astronomical measurement.

Acknowledgements

The funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

References

- Abazajian, K., *et al.*, 2003, AJ, 126, 2081 (DR1)
- Abazajian, K., *et al.*, 2005, AJ, 129, 1755 (DR3)

- Bahcall, J. N., 1968, ApJ, 153, 679
Bahcall, J. N., Peterson, B. A., Schmidt, M., 1966, ApJ, 145, 369
Bahcall, J. N., Spitzer, L., 1969, ApJ, 156, L63
Blanton, M. R., *et al.*, 2003, AJ, 125, 2276
Churchill, C., 2005, this volume
Fukugita, M., *et al.*, 1996, AJ, 111, 1748
Gunn, J. E., *et al.*, 1998, AJ, 116, 3040
Hall, P. B., *et al.*, 2002a, ApJS, 141, 267
Hall, P. B., *et al.*, 2002b, ApJ, 575, L51
Hobbs, L. M., *et al.*, 1993, ApJ, 411, 750
Hogg, D. W., *et al.*, 2001, AJ, 122, 2129
Inada, N., *et al.*, 2003, Nature, 426, 810
Ivezic, Z., 2005, AN, 325, 583
Khare, P., *et al.*, 2004, ApJ, 616, 86
Khare, P., *et al.*, 2005, this volume
Kulkarni, V. P., *et al.*, 2005a, ApJ, 618, 68
Kulkarni, V. P., *et al.*, 2005b, this volume
Menard, B., P eroux, C., 2003, A&A, 410, 33
Menou, K., *et al.*, 2001, ApJ, 561, 645
Nestor, D., Turnshek, D., Rao, S., 2003, ApJ, 595, L5
Pettini, M., West, K. A., 1982, ApJ, 260, 561P
Pier, J., *et al.*, 2003, AJ, 125, 1559
Prochaska, J. X., *et al.*, 2003, ApJ, 595, L9
Reichard, T. A., *et al.*, 2003a, AJ, 126, 2594
Reichard, T. A., *et al.*, 2003b, AJ, 125, 1711
Richards, G. T., 2001, ApJS, 133, 53
Richards, G. T., *et al.*, 2002a, AJ, 123, 2945
Richards, G. T., *et al.*, 2002b, AJ, 124, 1
Richards, G. T., *et al.*, 2003, AJ, 126, 1131
Richards, G. T., *et al.*, 2004, ApJS, 155, 257
Sandage, A., 1965, ApJ, 141, 1560
Schneider, D. P., *et al.*, 2003, AJ, 126, 2579 (DR1)
Schneider, D. P., *et al.*, 2005, AJ, in press (DR3)
Scranton, R., *et al.*, 2005, astro-ph-0504510
Sembach, K., *et al.*, 2000, ApJ, 538, L31
Smith, J. A., *et al.*, 2002, AJ, 123, 2121
Stoughton, C., *et al.*, 2002, AJ, 123, 485
Tolea, A., Krolik, J. H., Tsvetanov, Z., *et al.*, 2002, ApJ, 578, 31
Tytler, D., Fan, X.-M., 1992, ApJS, 79, 1
Vanden Berk, D. E., *et al.*, 1996, ApJ, 469, 78
Vanden Berk, D. E., *et al.*, 1998, AJ, 122, 355
Vanden Berk, D. E., *et al.*, 2005, AJ, 129, 2047
Wakker, B., *et al.*, 2003, ApJS, 146, 1
Weymann, R., *et al.*, 1979, ApJ, 234, 33
York, D. G., *et al.*, 1991, MNRAS, 250, 24
York, D. G., *et al.*, 2000, AJ, 120, 1579
York, D. G., *et al.*, 2001, AAS, 198, 7804